UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

OPTIMAL STRUCTURE PATTERNS FOR RESILIENT CORPORATE NETWORKS

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE

By

TAOUFIK ZORMATI
Norman, Oklahoma
2019

OPTIMAL STRUCTURE PATTERNS FOR RESILIENT CORPORATE NETWORKS

A THESIS APPROVED FOR THE
SCHOOL OF INDUSTRIAL AND SYSTEMS ENGINEERING

BY THE COMMITTEE CONSISTING OF

Dr Andrés D. González, Chair

Dr. Kash A. Barker

Dr. Charles D. Nicholson

Acknowledgments

Firstly, I would like to thank my advisor Dr. Andrés González for all his knowledge and advice which was determinant in the completion of this work. I really appreciate all he provides me.

Secondly, I am grateful to Dr. Toby Li for all his continuous help and advice.

My kind regards to Dr. Barker and Dr. Nicholson for being part of my committee.

Finally, I would like to thank my family for all the support they gave me despite the distance and particularly my wife for all her patience and encouragement.

Table of contents

List of figures

List of tables

I. Abstract

Corporate networks are sets of companies maintaining formal relations between them in the form of contracts (INSEE, 2016). This type of network is generally affected when unexpected events happen, such as financial crises or other disasters. The notion of corporate network resilience appears, and it is a field increasingly studied. This resilience is defined as the ability to adjust the activity of the system to retain its basic functionality when errors, failures, and environmental changes occur (Gao, Barzet & Barabási, 2016). It is critical for firms that are members of these corporate network to increase their resilience and get prepared for future unexpected events. Otherwise, being unprepared could lead to bankruptcy. We have seen in the past years the consequences of financial crashes on companies. The resilience of a firm in a corporate network is impacted by different elements such as the properties of the network, the size, and the structure. In this thesis, we focus on the impact of the structure of the network of a company on its resilience. We propose a way to analyze empirical data, in order to find the most resilient structure for different problems. We illustrate our methodology with a case study of a network of companies in the oil-gas industry and provide advices to obtain a resilient structure for a company to survive against downturns, economical or technological.

Keywords- Corporate networks, resilience, oil-gas industry, network structure

II. Introduction

The corporate networks are sets of companies maintaining formal relations between them in the form of contracts (INSEE, 2016). That type of network is directly involved when unexpected events happen, such as economical or technological crises and is put to the test. In this case, we focus on the concept of network robustness. Network robustness, which has been one of the most active topics in complex networks theory (Cahen & Halvin, 2010), refers to the structural resilience of a network to external perturbations. The robustness of this type of network has a huge impact on its firms. Companies are trying to be more resilient to different shocks that could make them go bankrupt. In fact, network recession, which is defined as a significant decline in economic activity, is difficult to forecast. These recessions happen often after economic shocks and can make firms fail. In order to prevent these disasters, it is important to be prepared and have an efficient structure because network structure highly matters for system resilience (Anand et al., 2011). Depending on the network and the type of crisis, the evolution can be entirely different for firms. In fact, companies can go bankrupt if they do not have an adapted network structure, or they can survive and grow if they are well prepared and adaptative. The study on the network robustness in the past years focuses not only on theoretical interests, but also on practical applications to design more resilient structures against random breakdowns or intentional attacks (Valente et al., 2004; Tanizawa et al., 2005; Schneider et al., 2011). In the same vision, we propose a way to find resilience patterns when working on a database including history of any kind of corporate networks, and to provide elements to companies trying to join the network in order to make them as resilient as possible.

In this thesis, we look firstly at the state of the art in the domain. We talk about the works concerning corporate networks in general in order to understand the ecosystem of companies whom we want to increase the resilience. Then we explore previous research on resilience, consequences of the damages, and different aspects of firm's resilience. It is crucial to understand the behavior of resilience and its properties to be able to improve it. Also, the impact of the structure of a network on resilience is an important point. It highly impacts the resilience and is one aspect that can be handled. In anticipation for a disaster, the structure is one of the first points that must be adapted. The lack of anticipation and preparation of the structure could easily lead firms to bankruptcy. We talk more about this aspect in the third and final part of the literature review.

Secondly, we present the methodology where we offer to find patterns to adapt the structure and increase resilience of all types of firms composing corporate networks. We present all the data needed, each of the specific aspects important to find the best structure, and all the steps to find it. We present how a firm can structure its network of partners to increase resilience and survive against downturns, working on the properties of its partners and the relationships with them.

Finally, we apply the methodology to our problem. In our case study, we work on a group of drillers and their network of client operators in the oil-gas industry in the United States. Drillers and operators form an oil corporate network. The nodes of our network are companies and links are contracts. To reach our main goal, which is improving the resilience of corporate networks in case of unexpected events, and providing the most adapted structure, we explore the 17 years of data available, and we present the characteristics of successful firms, and the path to follow to increase resilience. We also mention the properties that have to be avoided.

III. Literature review

1. Corporate networks

Networks, or graphs, are now widely used in economic and financial literature, as they represent a natural way to study connections and systemic effects. Under corporate networks we consider all networks that describe interactions between companies (Squartini, 2013). Various company interactions can be described by networks, that make them an efficient tool to study all type of events. To understand the ecosystem of companies, these interactions cannot be seen in isolation (Jeude et al., 2019). We consider all the aspects of the companies, to provide the most efficient answer to different issues. All the elements of that kind of network work together from an economic point of view. As a consequence, the global state of the network impacts all the companies forming the structure. When damage happen, most of the firms are involved and impacted. The consequences of network disruption have become increasingly severe (Sterbenz et al., 2010), leading to bankruptcy and failures. As a consequence, firms invest in having a resilient structure.

2. Network resilience

Resilience is a system's ability to adjust its activity to retain its basic functionality when errors, failures, and environmental changes occur (Gao, Barzet & Barabási, 2016). It is a defining property of many complex systems (Gao, Barzet & Barabási, 2016). Different aspects of resilience exist, some are inherent to the network and some are inherent to the firm. We consider four elements of community resilience: anticipation, reduced vulnerability, response, and recovery. Comparing local inherent resilience to formal corporate resilience enables the identification of their various strengths and weaknesses (Colten et al., 2012). In this research, we focus on how a firm that is trying to be resilient in a corporate network can structure its network of partners. Networks properties and structure affect the resilience of growth to economic shocks (Kharrazi, Rovenshaya & Fath, 2017). This is part of anticipation to a potential event. Networks are composed of two basic elements: a set of nodes and an arrangement of ties among the nodes. Network change, at its most basic level, will involve the modification of one or both of these elements (Hernandez & Menon, 2018). In corporate networks, if we consider the nodes as companies, and the ties as different types of contracts, the modification of the structure can be link removal (contract) or node removal (bankruptcy of a firm). Node deletion will affect network's evolution (Anand et al., 2011). The robustness of a network is the capacity to have an acceptable evolution even when important nodes are removed. Also, network robustness under removal of nodes (or links) depends on the connectivity patterns of networks (Min et al., 2014). The bankruptcy of one firm can take important proportion depending on its position in the network. Previous studies have shown that a small fraction of overloaded nodes extremely accelerates the propagation of failures (Yong-Hyuk & Yong-Sung, 2015). The events leading to loss of resilience-from cascading failures in technological are rarely predictable and are often irreversible (Gao, Barzet & Barabási, 2016). That is why having an adaptative structure can prevent damages and increase the robustness (i.e., the capability of surviving intentional and/or random failures) (Quattrociochi, Caldarelli & Scala, 2014).

3. Network structure

Different studies show that the topology of the networks influences the resilience to shocks. Network structure is twofold important, because first it determines the reaction of the network after an unexpected event but also, its evolution could even indicate early-warning signals of a crisis (Missaoui et al., 2003). Network structure can provide indication on the future events. Different parameters play their role, whose network size. Network size has an impact on resilience, if the size increases the resilience increases (Nagaishi &Takemoto, 2018). It has to be considered. Depending on our network, several types of topology can be of interest. For example, the kind of structure of the Darknet, which is considered as much more resilient than the Internet (used as a benchmark for comparison at a descriptive level) to random failures, targeted attacks, and cascade failures (De Domenico et al., 2017). It is characterized by a nonhomogeneous distribution of connections, typical of scale-free networks; very short path lengths and high clustering, typical of small-world networks; and lack of a core of highly connected nodes (De Domenico et al., 2017). Since our goal is to realize a study of corporate network resilience, a topic which has not been highly studied, we explore how network structure impacts the evolution and durability of firms through unexpected events. Our contribution is to provide an empirical way to study the resilience of any corporate network, and as a result to provide advices to firms joining the market.

IV. Proposed methodology

In this study, we are working on corporate networks. “A corporate network is a set of enterprises maintaining formal relations between them in the form of contracts between the business units” (INSEE, 2016). In other words, corporate network represents the relations between different companies. The goal of this methodology is to provide a general method to analyze and determine the most important properties of a given network to increase the resilience of a firm that is going to join or is already part of this specific market. Since we are leading an empirical study, based on the past years of the network, we need a database including several pieces of information.

We have seen in the previous part the impact of network structure on the resilience. The elements that make this structure at its most basic level are the partners of a firm: number, number of contracts shared, properties of the partners, the financial amount of the contracts, the duration of the contracts and others. In this study, to reach our goal, the type of information we use about the companies are:

- name,
- date of the first contract,
- date of last contract,
- duration of the company (total time when the company worked, if not provided we use calculate the difference between the last and the first contract),
- list of contracts of the company,
- list of the partners, in order to analyze their properties.

We are not taking into account the amount of the contracts. We are considering all the contract with the same value, but further work may include this variable to determine its impact on the resilience. Our global approach is divided into three main parts. First, we realize a preliminary analysis in order to understand the context of the problem, the general properties, and find the important variables. With that overview, we can analyze the network deeper. In the second step, we look for the main links and correlations between our variables. At this step, we can determine the variables that highly matters to increase resilience. This part allows us to realize the final one. Working with the most important variables on each particular problem, we can apply our selection of individuals and clusters in order to find the best adapted structure and its properties, but also the types of structure to avoid. Finally, we can show the most interesting properties for the specific problem.

1. Preliminary analyses

The first step here is to determine what makes a resilient company. We consider in this study that the duration of a company is a crucial aspect of resilience. Having worked for a long time through different episodes is a symbol of success. Avoiding bankruptcy and continuously working is our metric of resilience. Therefore, the duration of the companies is our variable of interest in this analysis. In this first step, if we have at least two types of companies working in the domain, we realize a statistical test in order to compare them. It is important to know if we can consider the types of firms as similar for the distinct variables. We compare the means of each group with an analysis of variances (ANOVA), considering

k types of companies. We want to know the differences among group means in the sample. The hypotheses are the following.

Hypotheses:

- $H0 : \mu1 = \mu2 = \mu3 \ldots = \mu k$
- $H1 : Means\ are\ not\ all\ equal$

We have two hypotheses, the first one, the null hypothesis is that for the observed variable the means are considered as equal, the second one is that at least one of the means is considered different than another. When performing ANOVA, we must check the assumptions:

- response variable residuals are normally distributed,
- variances of populations are equal,
- responses for a given group are independent and identically distributed normal random variables.

If all the requirements are met, it is possible to perform the ANOVA and realize the comparison of variances. Otherwise, it is possible to realize a non-parametric test such as the Kruskal-Wallis analysis of variance. After the test we reject or not the null hypothesis. If we reject it, to know exactly which groups are different we can run a test such as the Least Significant Difference test (LSD). If we are working on only two sets of companies, we can use the following figure to choose the test to perform.

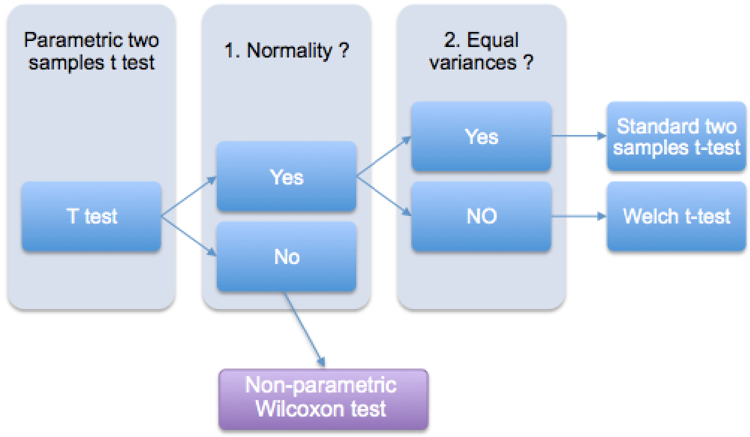

Figure [1]: Parametric two sample test

We apply this process to different variables. First, to the average duration of companies depending on their group in order to know if their means are considered as equal or different. This first test provides an understanding of the general duration behavior for each type of firm. Second, we do an analysis of variance on the average number of partners. Depending on the type of company, the properties can be different, and this directly determines the structure of the network. We also perform the process on the average number of contracts of each type of firms, to complete this preliminary analysis. These variables directly impact network structure and are considered significant. This first part gives an overview of the general behavior of the different types of companies and allows us to go deeper in the analysis.

2. Correlation studies

It is important to analyze the impact of the main variables on our variable of interest, the average duration of companies. We have previously observed two important variables for different types of firms. Theses variables determine the network structure at it most basic level, the average number of partners and contracts. In the network, it represents basic nodes and edges that form the general structure. Now, we add two other variables that may be significant. Firstly, the duration of the partners of a company. The general profile of the partners of a company (high or small resilience) may have an impact on the way the company work. Working with big companies may bring resilience to the structure. Secondly, we introduce the average betweenness. Considering that a company is part of a network, this is the number of times that a node is in the shortest path among two others. The shortest past from one node to another in a network is the path between two vertices such that the sum of the weights of its constituent edges is minimized. In our case all the edges have a weight of one. If this value is high, that means the node is highly connected, and has an important place in that network. In reality it can be the case for a big company. This variable can potentially have an impact on the network structure and the resilience. Being highly connected may increase the resilience.

We run correlation analyses between the average duration of companies and the four other variables to determine and classify their impact on the resilience. We apply this process to the different types of companies and obtain a list of the variable that impacts the average duration of each, in other words, the resilience.

3. Clustering and group creation

The previous step helped us to determine the important variables for our problem. We select the three highly correlated variables with our variable of interest in this part. To have an overview on the behavior of the variables, we start by clustering the data. We want to see how the three variables react when our variable of interest, the average duration of companies, changes. For this step, we choose to use the k-means clustering algorithm, for each type of company. We impose the value of $k = 2$ in order to obtain two groups, one with a higher duration and one with a lower duration, depending on the problem. The k-means algorithm works in that way: first it starts with a group of randomly selected centroids, which are used as the beginning points for every cluster. We impose two clusters in our case and each one has a centroid. Then the algorithm performs iterative calculation to optimize the position of the centroids, in order to minimize the distance between the points and the centroid inside of each cluster. We want to minimize the following value.

$$\text{objective function} \leftarrow J = \sum_{j=1}^{\overset{\text{number of clusters}}{k}} \sum_{i=1}^{\overset{\text{number of cases}}{n}} \underbrace{\left\| \overset{\text{case } i}{x_i^{(j)}} - \overset{\text{centroid for cluster } j}{c_j} \right\|^2}_{\text{Distance function}}$$

Figure [2]: K-means algorithm objective function

After performing the algorithm, we obtain two groups. We observe the general behavior of the three variables when we increase the average duration of companies. After this step, and depending on each problem, we choose to create four arbitrary groups depending on the duration of the firms. For example, in the first group the firms that worked less than one year; in the second, firms that worked between one and three years; in the third, firms that worked from three to five years; and in the last group firms that worked more than five years. We split the companies into the groups for each type of company to observe the values of the three other variables depending on their group. We check the number of companies in each one of the groups, to see the proportion that was successful, and the proportion that failed. Also, if we have a long period of data, we realize the same analysis and split the companies into groups depending on the age of firms to see if the period of apparition has an impact on the average duration. If it is the case, we cross the different tables. For each period of apparition, we create four groups based on the duration of companies to have a deeper vision. After this and in order to give indications to companies to increase their resilience, we choose to create two levels of each of the three variables we selected after the correlation analysis. We call the variables A, B and C, and the two levels are – and +. We place the fifty percent of the companies with the lowest values of the variable in the minus group and the fifty others in the plus group. We create a table with all the combinations between the three variables and their levels.

Table [1]: Combinations of variables

A	B	C
+	+	+
+	-	-
+	+	-
+	-	+
-	+	-
-	+	+
-	-	-
-	-	+

In the table, we include the number of companies for each combination, the average duration and the average age. We replicate it for each type of companies and based on this, we can find the groups that have the highest durations and the lowest ones and their properties. We obtain the properties of the resilient firms and of the ones not resilient. At this stage of the analysis, we know what structure has to be imitated and what structure has to be avoided. The results from this analysis give indications to companies on how to increase their resilience.

V. Computational example

1. Problem

In this case study, our particular network is composed of drillers (contractors) and operators in the oil-gas industry in the United States. We assume that our two companies are the nodes of the network and contracts are the edges between them. In our database, we have only contracts between operators and contractors but never between two entities of the same type. The network forms a bipartite graph.

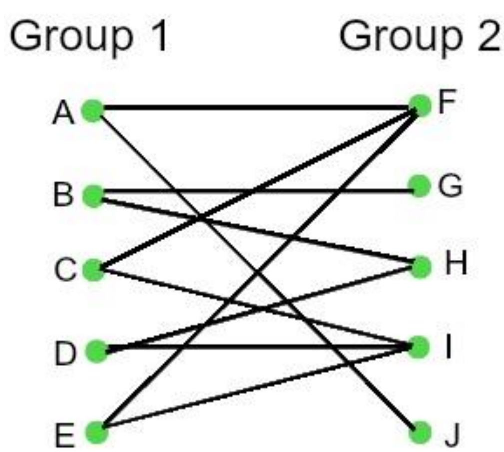

Figure [3]: bipartite graph

In graph theory, a bipartite graph is a graph whose vertices can be divided into two disjoint and independent sets U and V (for example on figure (3), group 1 and group 2); every edge connects a vertex in U to one in V. The graph does not contain any odd-length cycles which means that if two vertices are in the same group, they cannot be connected. A graph is bipartite if and only if it does not contain a cycle of odd length. This configuration considers two types of companies such a set of rental companies and a set of maintenance companies. Each rental company has its own apartments and call one or several companies to do maintenance. Rental companies or maintenance companies cannot work with companies from the same type, as concurrent. In our computational example, we work on a set of contractors and a set of client operators in the oil-gas industry. Many different structures can appear as seen on figure (3). For example, we can have one node linked to only one, or one linked to three. That means that a company can work with one partner or more, with any number of contracts. This makes the network structure. We are looking for the most resilient structure that can survive through different episodes, time, oil price downturn or any other crisis. In this study, to increase the resilience of companies, we use the previous data and analysis through the time. We are looking for companies that survived the longest time on a dataset starting in 1999 and ending in August 2016. In total, there are 6208 days, involving 856 operators and 139 contractors. We consider the companies at the same value and importance, which means having a contract with a big company is equivalent to having a contract with a small one. In reality, it can be different. Working with a huge company can bring resilience to the system due to its stability. We want to know how to structure the network of partners of a firm, to have the strongest resilience.

2. Hypotheses

We suppose in our problem that the type of partners of a company, and the relationships they share have an impact on the network resilience of the firms. Partners influence the resilience in the case of an unexpected downturn. In fact, knowing that if all the partners go bankrupt, the firm will go bankrupt. It is important to choose the right partners, the right number and the number of contracts to share. To display a successful company, we assume that the most important element is the duration of the firm which is the total time that a company has been working. If the firm survived for a long time through different crises, the network structure associated is considered as resilient.

3. Analysis

a. Preliminary analyses

In this first part of the analysis, we look for the general context of the problem. We consider a period of 6208 days starting in 1999 and ending in 2016. On this period, we look first to the duration of companies considering 139 contractors and 856 operators.

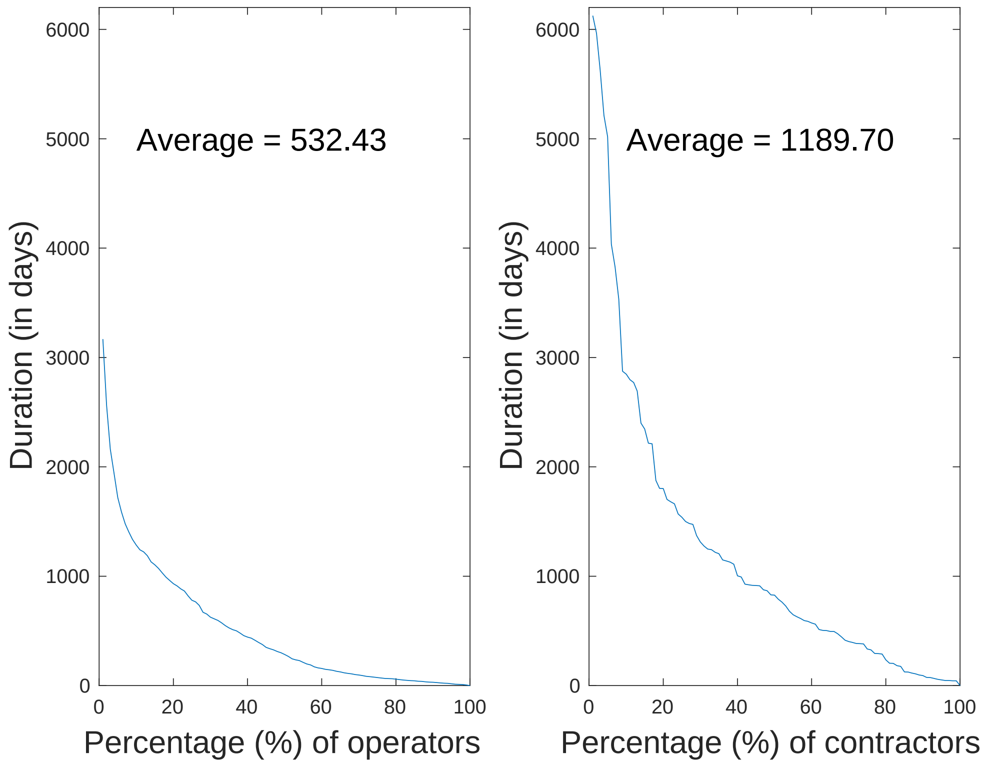

Figure [4]: Average duration of operators and contractors

The average duration of contractors is 1189.7 days, equivalent to 3.25 years. For operators, it is 532.4 days, so about 1.46 year. On average, contractors work more than 2.20 times longer than operators. The comparison of means can determine if the two types of entities can be considered as similar for the average duration. Using figure (1), we look for the test to perform in this two-sample case. Using this dataset, we consider that individuals are independent, and we check the two other assumptions. The normality assumption first:

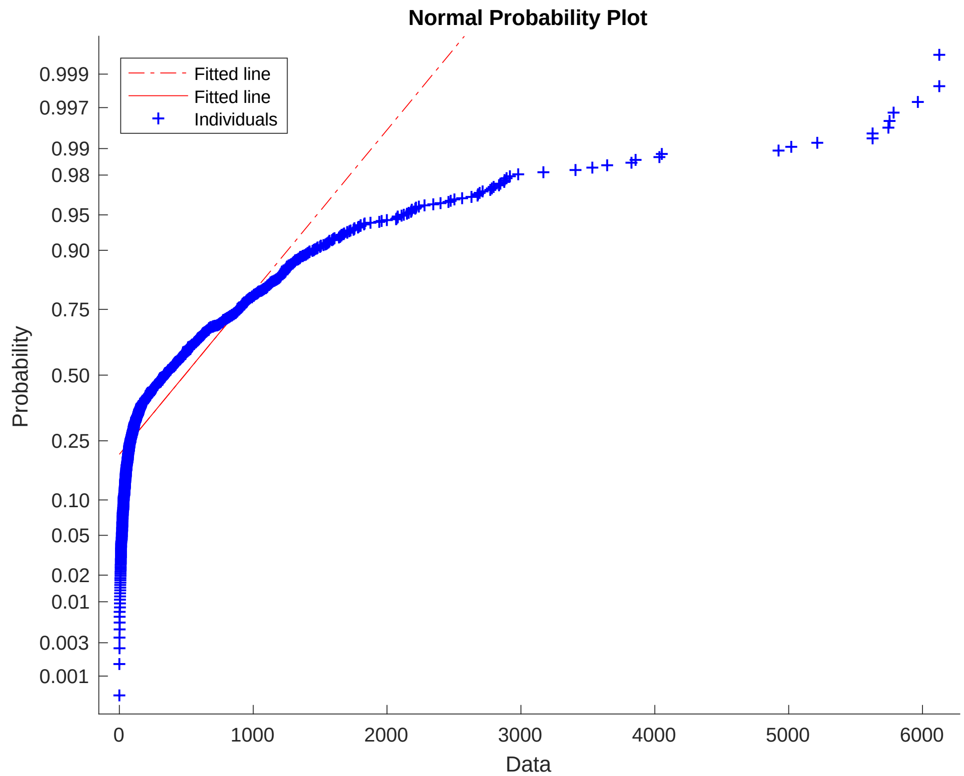

Figure [5]: Normality plot for duration of companies

The data does not follow a normal distribution. We decide to run the non-parametric Wilcoxon rank sum test to deal with two populations with two independent samples with different sample size. This test can be applied in the one factor ANOVA case as we have, when the normality assumptions do not apply. This test determines whether the medians of the groups are different. The hypotheses are:

- $H0: medians\ of\ contractors\ and\ operators\ are\ equal$
- $H1: medians\ of\ contractors\ and\ operators\ are\ not\ equal$

The result for this test is a p-value of 1.70e-13, so we reject the null hypothesis. We cannot assume that the medians of the two populations are equal. The total duration variable is different depending on the type of company. Drillers are more resilient on average as companies than operators knowing that they work more than twice longer. Since we are studying network properties, we must to introduce the notion of degree. The degree of a node is the number of edges connected to it. In this problem it can have two interpretations. If the nodes are companies, we can consider the degree as the number of contracts or as the number of partners. We analyze the number of partners for the two types of company first. The following plot is the average number of partners of each type of entity, day by day during the total period.

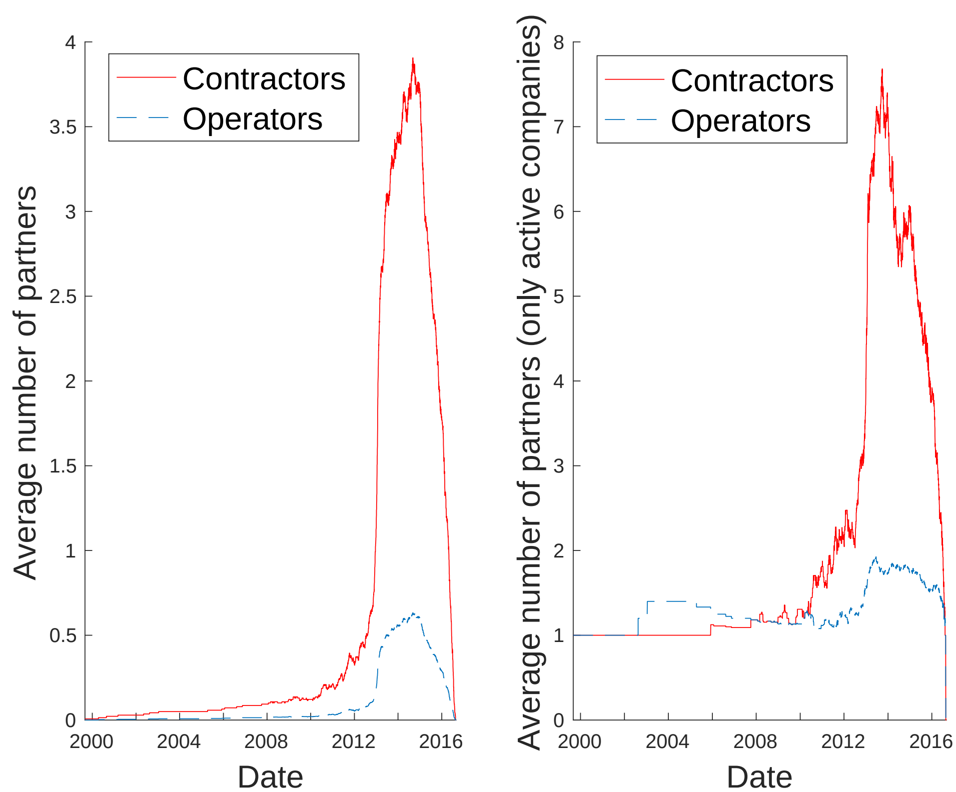

Figure [6]: Average number of partners of contractors and operators, all entities and only active entities

The average number of partners is higher for contractors than for operators, reaching almost 4 and only 0.70 respectively at its highest value on the plot 1 of the figure (6). This plot includes all the elements of each entity. The second plot of the figure (6) takes only into account the active entities (ones which are working the corresponding day). As a consequence, the average values are higher than the previous values because we are not taking into account the inactive entities (ones without contracts for the corresponding day). When new patterns appear, we see that the average number of partners of active operators is steady during almost all the period, and it has few differences depending on time, always between one and two. Operators on average work with a constant number of contractors. On the other hand, the average degree of active contractors has a big increase from day 2010 until 2014. This change in the behavior of a contractors signifies that contractors increased the number of operators they worked with when they reached a peak period. Contractors do have a different behavior compared to operators; they can increase their number of partners depending on the period. We want to know if the number of partners can be considered as equal for our two entities. As for the average duration of companies, we start by checking the assumptions to run the ANOVA. First, the normality assumption.

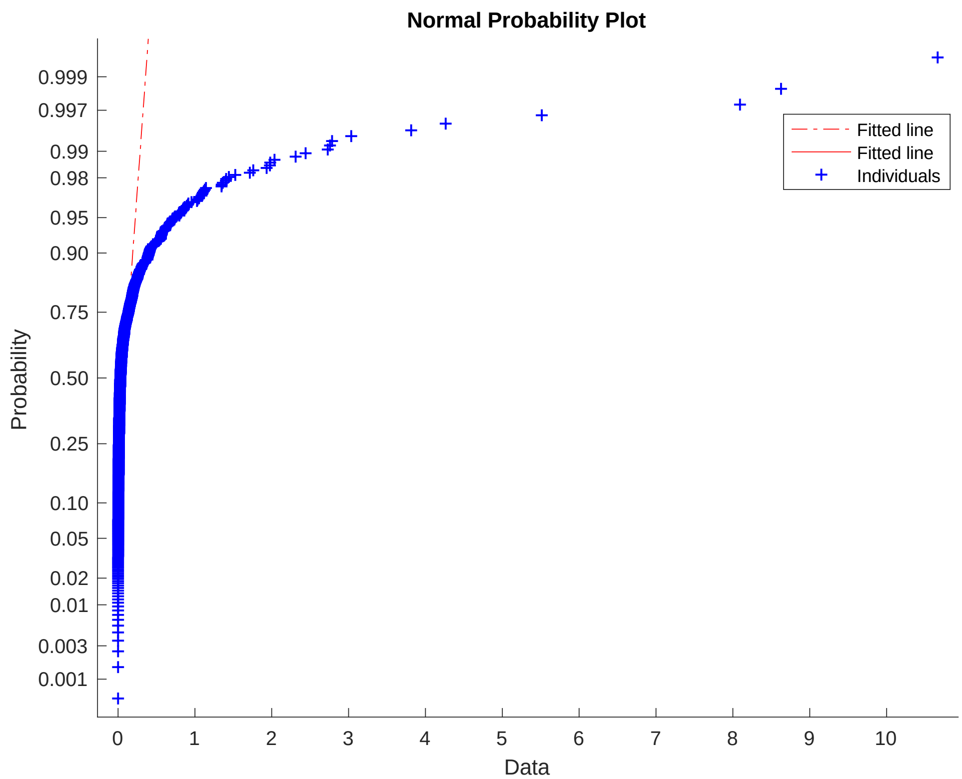

Figure [7]: Normality plot for number of partners

According to figure (7), the data does not follow a normal distribution. As a consequence, we run the non-parametric Wilcoxon rank sum test, having the same assumptions and the same hypotheses as for the duration of companies.

- $H0 : medians\ of\ contractors\ and\ operators\ are\ equal$
- $H1 : medians\ of\ contractors\ and\ operators\ are\ not\ equal$

We obtain a p-value of 2.67e-15 so we reject the null hypothesis. We cannot assume that the medians of the number of partners are equal for our two types of companies. We look now at the average number of contracts to complete the first analysis of the network. The following plot is the average number of contracts of each type of entity, day by day.

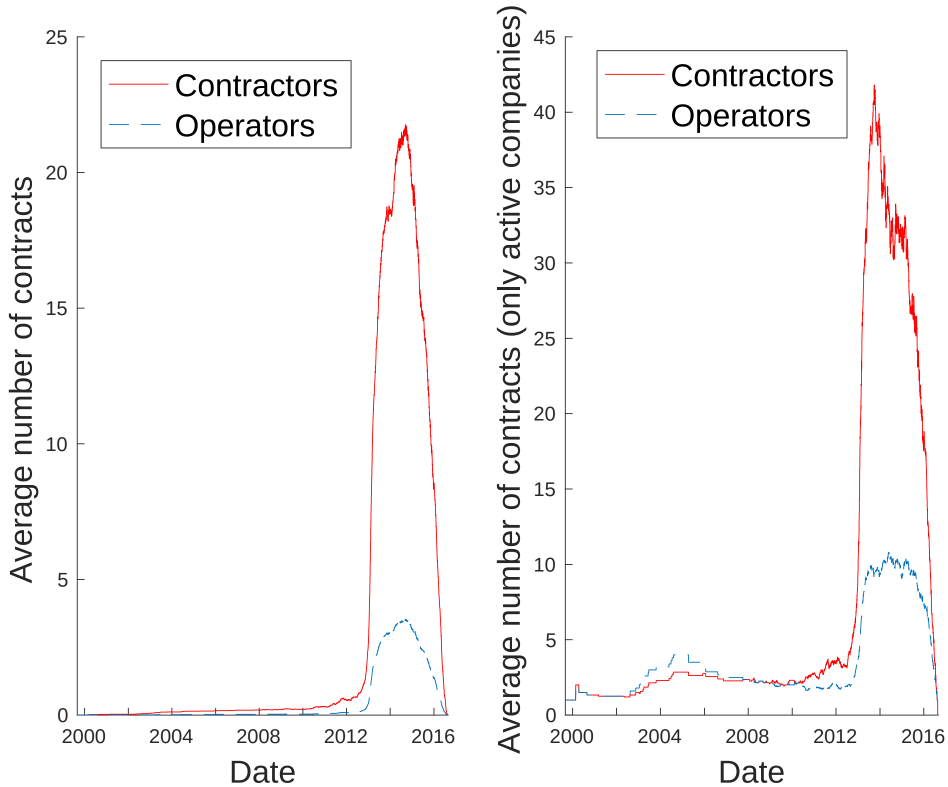

Figure [8]: Average number of contracts of contractors and operators, all entities and only active entities

On the first plot of figure (8), the maximum average degree is about 23 and 4, for contractors and operators respectively, but the two curves have the same shape. As we did previously, we use only the active entities on the second plot to calculate the average number of contracts. We can see that it increases around 2013 for both types of companies. Adding this to the information provided on figure (6), operators increase their number of contracts but not the number of companies. They increased the number of contracts with companies they were working with. According to figure (6) and figure (8), we can see that contractors increased their number of partners and their number of contracts. Instead of increasing the number of contracts with their partners, contractors created new links with new partners. Knowing these differences between the two types of firm, we check the equivalence of means for the number of partners variable. The normality assumption first.

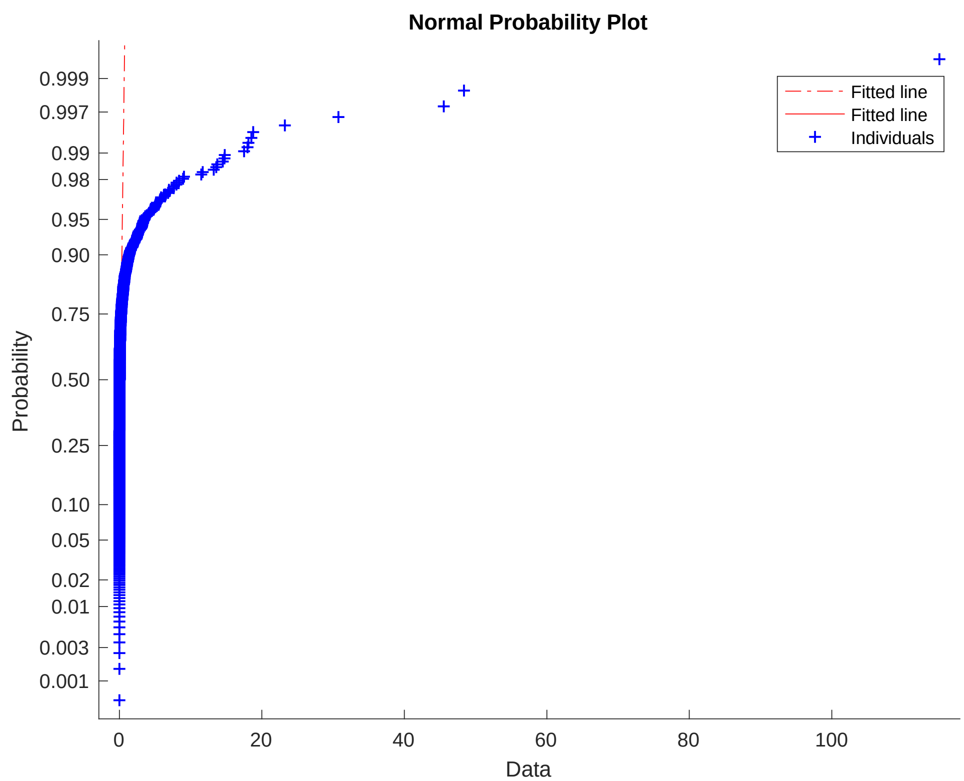

Figure [9]: Normality plot for the number of contracts

The normality assumption does not hold. As for the average number of partners, we realize the non-parametric Wilcoxon rank sum test with the same assumptions. The hypotheses are the same.

- $H0 : medians\ of\ contractors\ and\ operators\ are\ equal$
- $H1 : medians\ of\ contractors\ and\ operators\ are\ not\ equal$

We obtain a p-value of 3.68e-13 thereby we reject the null hypothesis. We cannot assume that the medians of the number of contracts are not equal for contractors and operators.

After this first study of the most basic variables, we see that the contractors and operators are two clearly different types of companies. For each one of these, we realize a deeper study in the following step and look for the impact of different variables on the average duration of companies.

b. Correlation studies

In this part, we look for the impact of the variables on the average duration of the two types of company. We run correlation analyses between our variable of interest and four other variables. These variables are the number of partners of a company and its number of contracts, we have analyzed these variables in the previous part, and they make the structure of the network of a company at its most basic level. Then, we add two new variables that may be important, the duration of the partners of a company which can be determinant to make a company more resilient and the average betweenness to analyze the impact of the connectivity of a firm. First, we start with a correlation analysis between the average duration of entities and the number of partners.

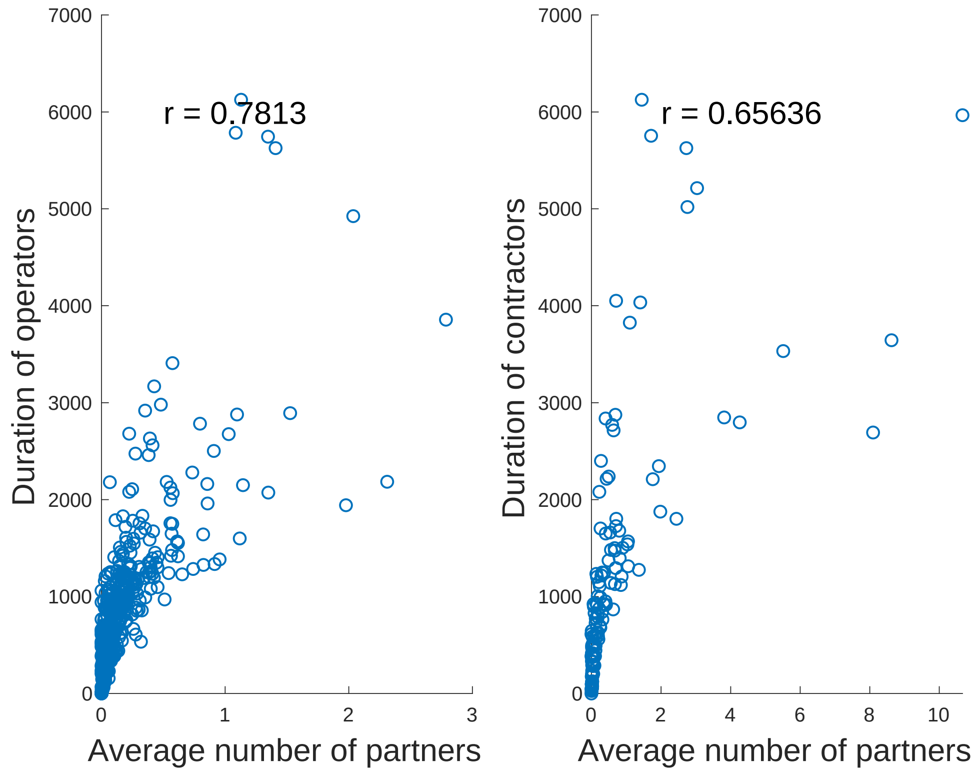

Figure [10]: Average duration of entities versus average number of partners

The coefficient of correlation is $r = 0.65$ for contractors. This implies that the average number of partners of a contractor has an impact on its duration. Contractors should work with as many as possible operators to increase their average duration. For operators the coefficient of correlation is $r = 0.78$, which means than the average number of partners and duration of an operator are highly correlated. The higher the average number of partners for operators, the longer it will stand. Operators should work with as many drillers as possible. In fact, we have seen in figure (6) that the global average of partners for operators is always between one and two. Getting new connections may be complicated in reality for operators. Two conclude with this first correlation, the two types of companies have their duration correlated to the number of partners and more particularly operators. After this, we look at the number of contracts to see its impact on the average duration.

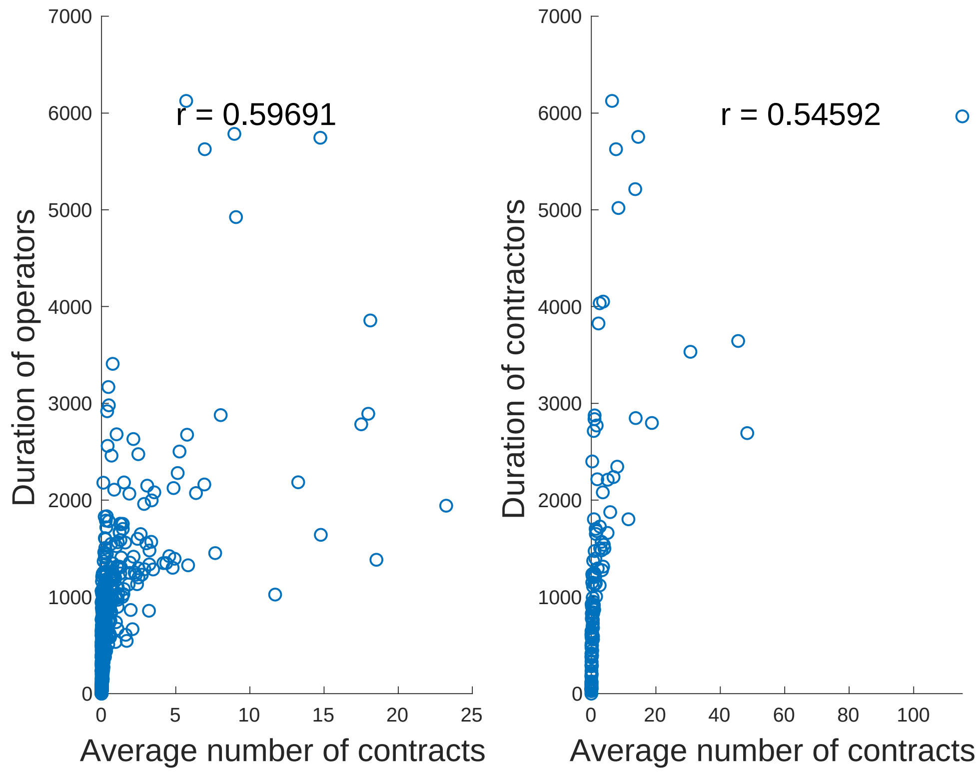

Figure [11]: Duration of entities versus average number of contracts

The coefficient of correlation is $r = 0.54$ for contractors which is less important than the previous one. The average number of contracts has less impact on the average duration than the average number of partners for contractors, but it matters. Operators behave similarly. The number of contracts and the number of partners have an impact on the average duration of the companies, but the quality of the partners should also bring resilience to the structure and make a difference. The partners are called neighbors, in reference to the network. We realize the correlation analysis between the duration of a firm and the average duration of its partners.

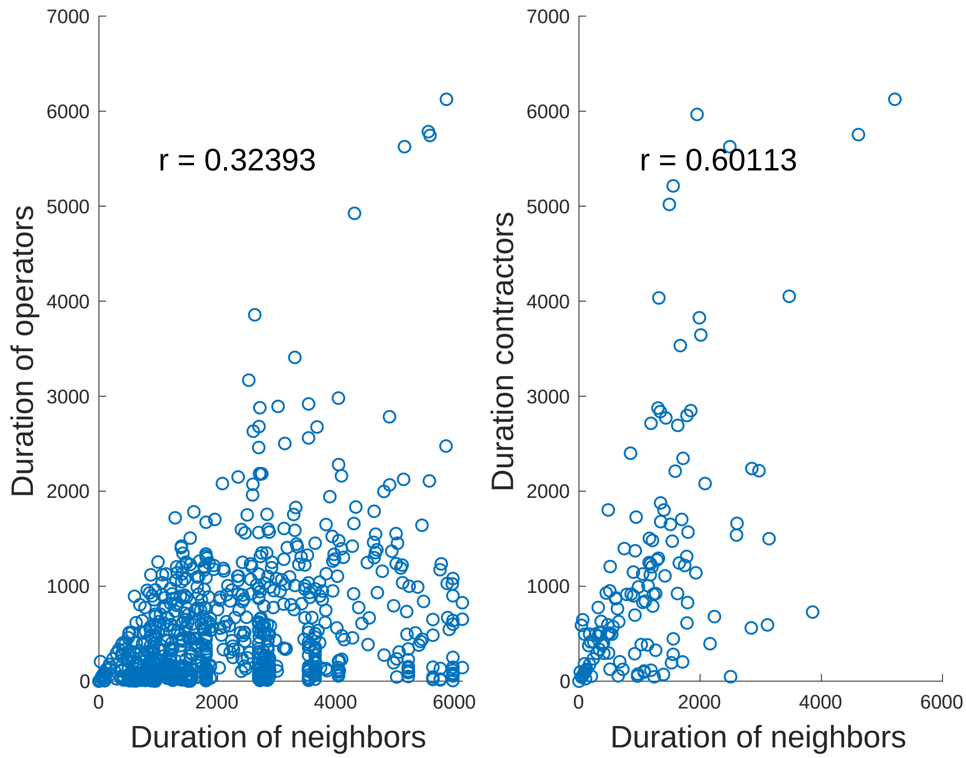

Figure [12]: Duration of entities versus average duration of neighbors

We have a correlation of $r = 0.60$ among duration of contractors and their neighbors. The duration of the partners can have an influence on the duration of a contractor. Contractors should work with long duration operators to increase their average duration. For operators, the relation between the two variables is almost inexistent ($r = 0.32$). It implies that operators should not necessarily look at the average duration of their partners. After this step, we look at our final variable, the average betweenness of a node. This is the number of times that a node is in the shortest path among two others. If this value is high, the node is highly connected and has an important place in that network. Big companies can have a high betweenness in real life.

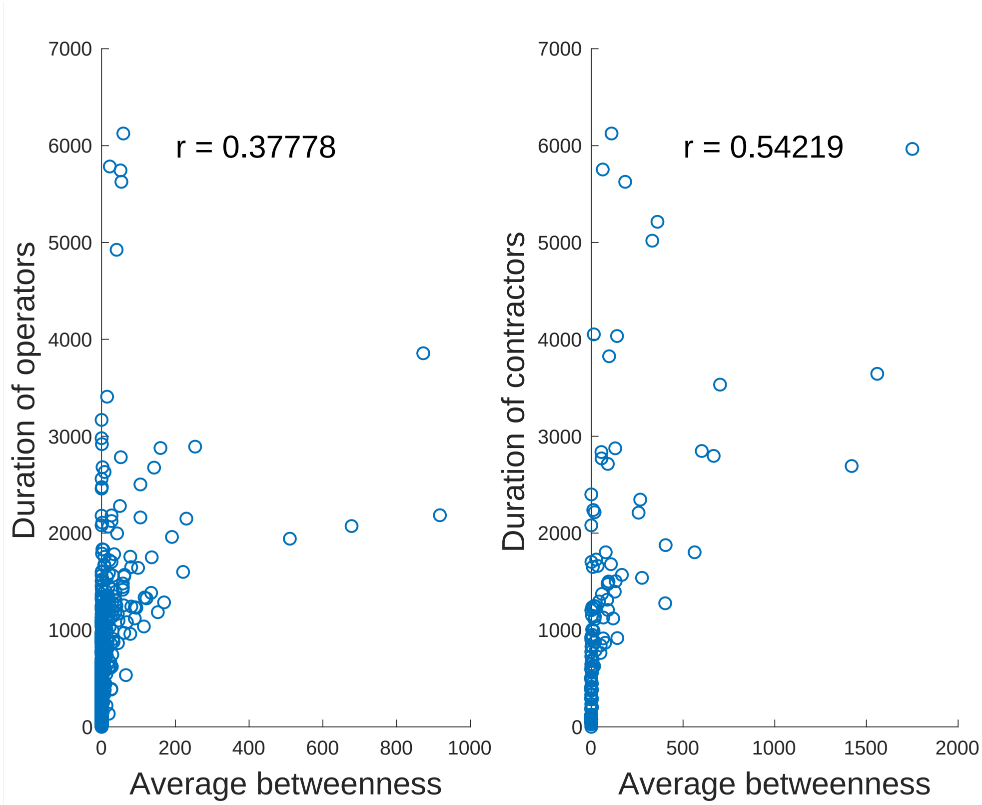

Figure [13]: Duration of entities versus average betweenness

Duration of contractors and average betweenness are connected ($r = 0.54$). It is better when working as contractor to be highly connected. In case of downturn, having multiple partners can have a positive impact, knowing that if your partners go bankrupt you will go bankrupt. The more partners you have, the less risk you have. For operators, the coefficient of correlation is smaller ($r = 0.38$), which means the duration of an operator does not depend on the betweenness.

To conclude this second part of the analysis, the duration of our two types of companies is highly correlated with the average number of partners, the correlation is smaller with the average number of contracts but exists. Operators are not correlated to the duration of the neighbors and the average betweenness, when contractors are connected to these variables with medium values. We choose to continue our study on three of the four variables, the number of partners, the number of contracts, and the duration of the partners. In this case they are the most correlated variables to the average duration of the companies of the two types.

c. Clustering and group creation

Before realizing the clustering and the selection of the best fitted structure in this case, it can be useful to have complementary information about the database. So far, we have dealt with the properties of the different entities, knowing that they are about 856 operators and 139 contractors. It can be of interest to see how many are active each day and how many are existing; which means that these companies do not have a contract this particular day, but they will have later; to improve our comprehension of the case study.

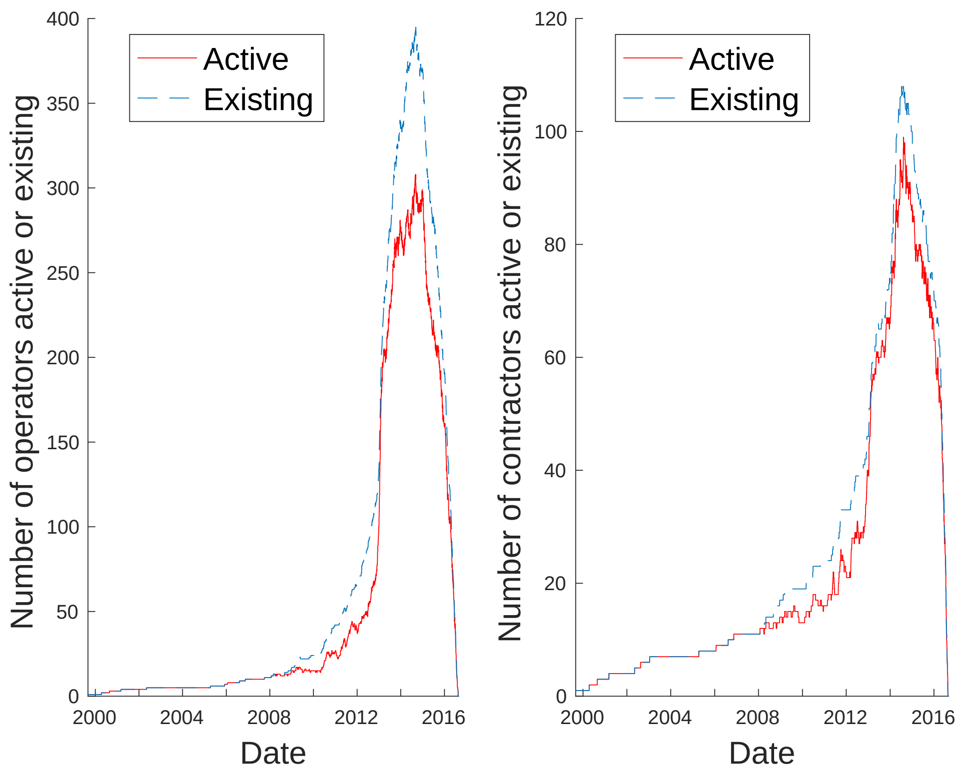

Figure [14]: Number of entities active or existing

For operators, the number of active and existing entities increases until 2013, but only about the half of the total number are existing for the peak period (400 out of 856), and about 25% of the existing operators are not active. The state of contractors seems to be more stable. The majority of them are existing around 2013, and almost all the existing ones are working. Another point to analyze is the average number of contracts per company and the price of oil for all the period.

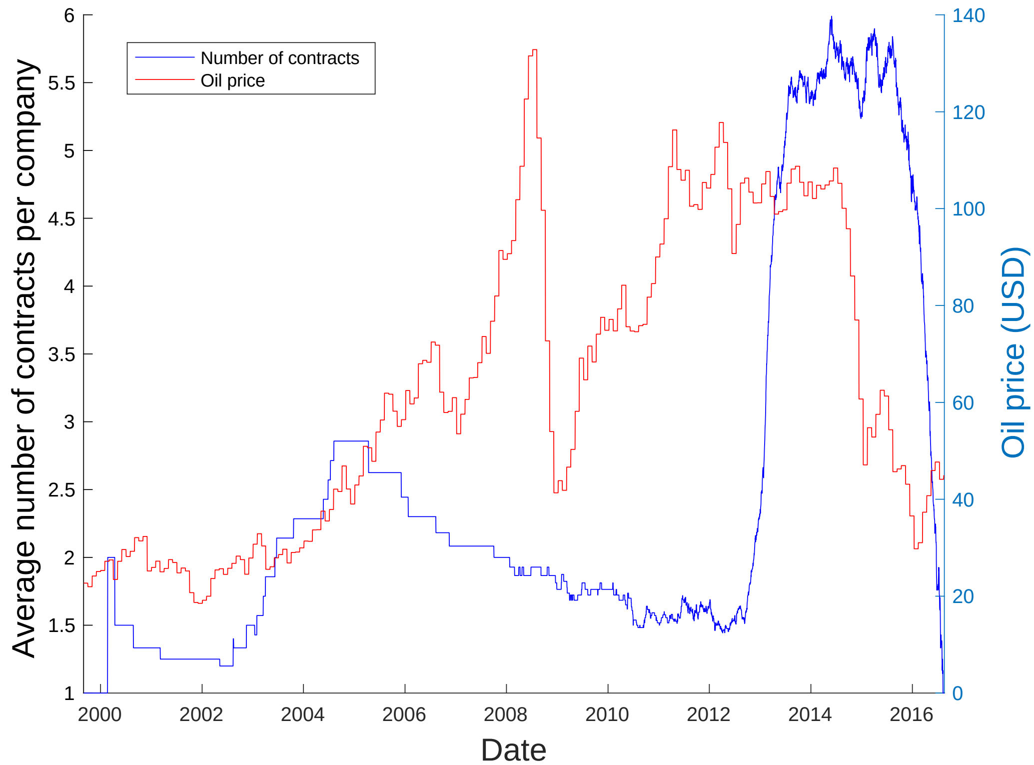

Figure [15]: Average number of contracts per company and price of oil

We have the average number of contracts per company, which represents the number of contracts per operator or contractor per day. We have on the same plot the price of oil per barrel day by day. We have seen that the previous plots (figures (6), (8)) depending on time were quite steady or increased slowly from 1999 until 2010. Then they reached a peak period around 2013. If we analyze the price of oil during these periods, we see an increase on both lines after 2002. However, we do not have clarity for the following periods due probably to another type of shock or to the end of the dataset. After this general part of the study, we make a selection between the companies to find resilient structure patterns. We are looking for the companies that survived the longest time and analyze their properties. In this part, we consider four variables which will help us to analyze our companies. Firstly, the duration of the companies is the time when companies stand, and it is the most valuable variable. Secondly, the average number of partners of the companies. Thirdly the average number of contracts per day and finally the duration of the partners. We choose these variables after the correlation analysis part. Knowing that we have two types of entities, we assume with different properties, we analyze each one of them. First of all, because we are considering the duration of the company as the most significant variable for the resilience of its network, it is interesting to realize a clustering and check the values of the other variables depending on the group. We do a k-means clustering with two groups and obtain the following tables.

Table [2]: Cluster of operators

Operators	Average duration (in days)	Average duration neighbors (in days)	Average number of partners per day	Average number of contracts per day
Group 1	1561.60	3092.70	1.40	7.60
Group 2	141.10	1215.10	0.40	0.70

Table [3]: Cluster of contractors

Contractors	Average duration (in days)	Average duration neighbors (in days)	Average number of partners per day	Average number of contracts per day
Group 1	839.80	1564.70	1.50	4.10
Group 2	604.90	1400.60	1.00	3.90

We can see on both tables (tables (2) and (3)) that for the group with the highest duration, all the variables are higher than the other group. This means for this dataset, having worked for a long time is linked with having neighbors with high duration, and having high number of contracts and partners per day. The three variables are positively correlated to the average duration. This first step shows the relation among all the variables and our variable of interest. To confirm the hypothesis that the higher is the duration, the higher are the other variables, we separate the entities of each type in four arbitrary groups with different durations. The first group includes the companies that worked less than a year, the second one companies that worked between a year and a thousand days, the third one between a thousand days and fifteen hundred days. The fourth group are the companies that worked more than fifteen hundred days. The results are on tables (4) and (6), operators and contractors respectively.

Table [4]: Groups of operators depending on duration

Operators	Average duration neighbors (in days)	Average number of partners per day	Average number of contracts per day	Number of companies	Percentage (%)
Group 1	1996.20	1.00	1.10	475	55.50
Group 2	2325.70	0.80	1.90	228	26.70
Group 3	2951.20	1.40	7.30	94	11.00
Group 4	3536.20	1.80	10.50	59	6.80

Operators have the property that each variable is increasing when the duration increases. We can see that only 6.8% of the total amount of operators survived more than 1500 days which is a really small part, and more than 55% survived less than one year. The higher the duration of work, the lower the proportion of operators. A correlation analysis can help to see the behaviors. First between duration of operators and duration of neighbors (plot 1 of figure (16)).

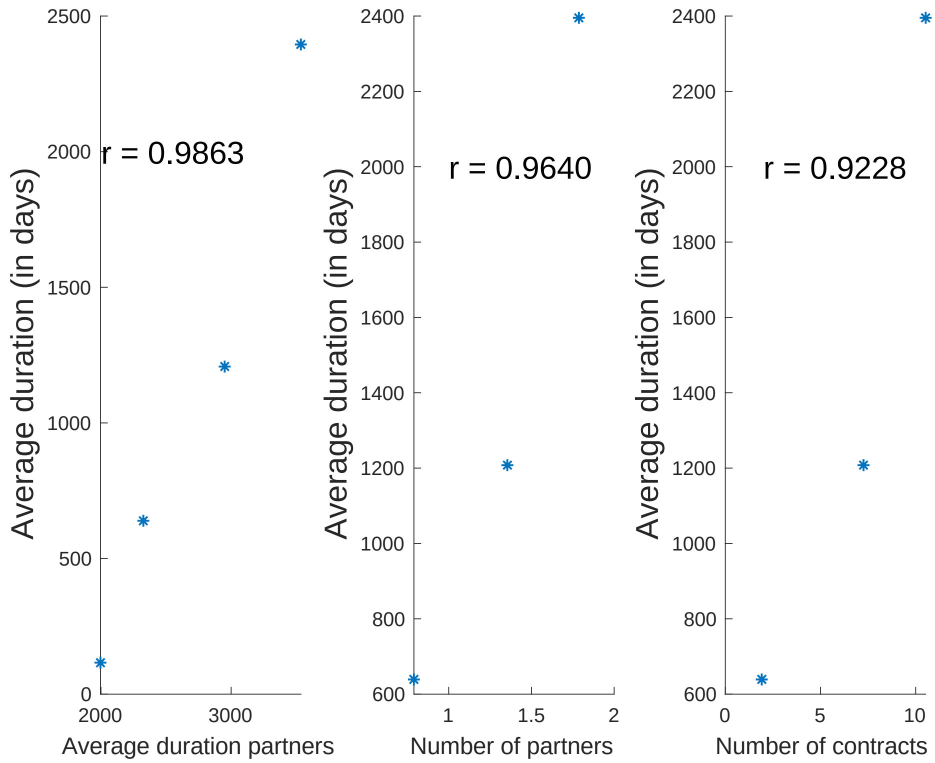

Figure [16]: Groups of operators versus duration of partners, number of contracts and number of partners

The first coefficient of correlation is $\boldsymbol{r = 0.99}$ which implies that duration of an operator and duration of its partners are highly correlated. We have the same high correlation between duration of the company and the number of contracts and the number of partners (plots 3 and 2 of figure (16) respectively). According to the different values we found, the duration of operators is highly influenced by the duration of the neighbors, the numbers of contracts, and the number of partners. We repeat these steps for the drillers. We have the following table.

Table [5]: Groups of contractors depending on duration

Contractors	Average duration neighbors (in days)	Average number of partners per day	Average number of contracts per day	Number of companies	Percentage (%)
Group 1	635.90	0.90	1.00	36	25.90
Group 2	906.80	1.30	2.10	48	34.50
Group 3	1256.20	2.70	6.80	19	13.70
Group 4	1955.00	4.20	20.80	36	25.90

We observe exactly the same patterns for contractors concerning our hypothesis. Also, when our main variable increase, the three other variables do as well. We check on the following graphs to confirm this pattern. Whereas, the repartition of the companies into the groups is different than for operators. Only about 26% of contractors survived less than a year (about 55% for operators) and about 26% survived more than 1500 days (about only 6.8% for operators). Contractors are more successful on average than operators.

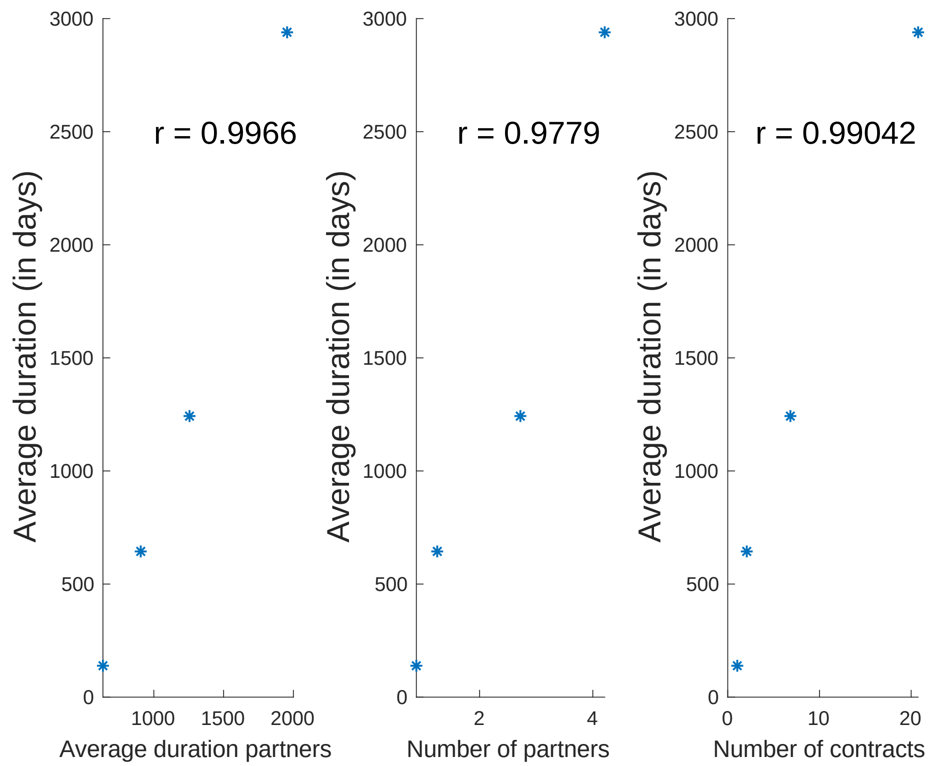

Figure [17]: Groups of contractors versus duration of partners, number of contracts and number of partners

All the coefficients of correlation are high, and we have the same properties that we found for operators. We can confirm that the duration of partners, the number of contracts, and the duration of partners increase when the duration of contractors increases. After this first step, we introduce a new variable, the age of companies. We have data of about 16 years, so we can assume that the time when a company appeared can influence the path it follows, and this time can be linked to resilience. We are now considering the age of the companies, and we realize another analysis. To create groups, we can find that they are three different periods on figure (18), which is the average number of contracts per companies. As a consequence, we choose to divide our companies into three groups.

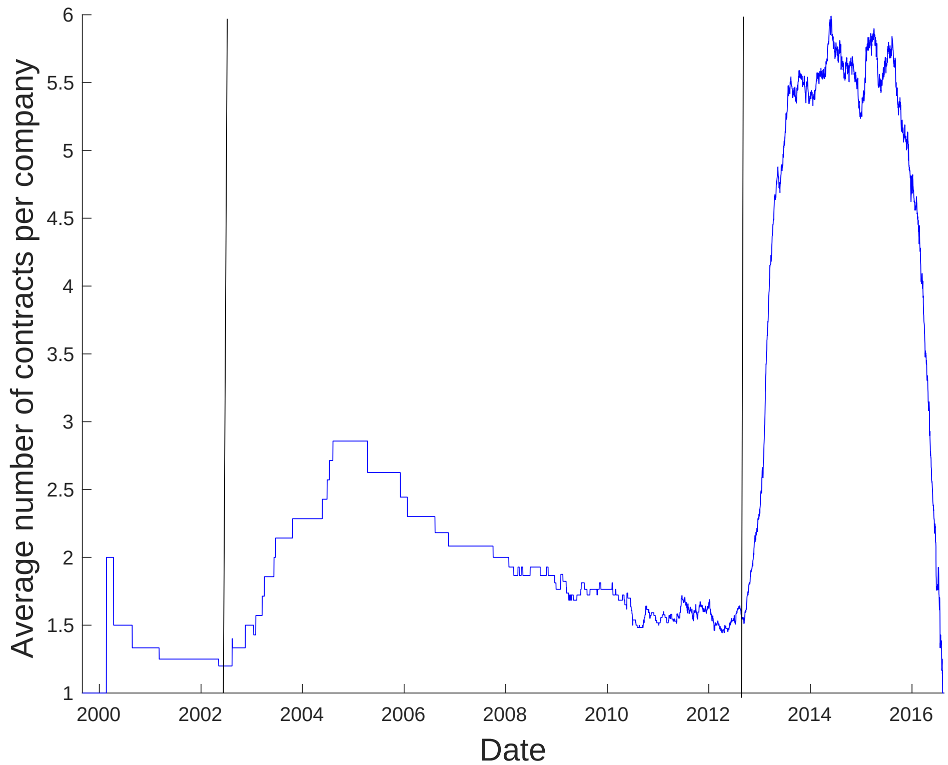

Figure [18]: Average number of contracts per company

First, the companies introduced the market from 1999 to February 2005. The second group is from the period of March 2005 to October 2012, and the last group is composed of the youngest companies, from November 2012 to August 2016. We analyze the companies considering the period when they appeared. With those three groups we get the following tables.

Table [6]: Properties of operators depending on age

Operators	Average duration (in days)	Average duration of neighbors (in days)	Average number of partners per day	Average number of contracts per day	Number of compani es	Percen tage (%)
Period 1	5641.20	5288.56	1.58	10.08	5	0.60
Period 2	1448.21	3375.85	1.46	7.74	125	14.60
Period 3	349.87	2096.19	0.94	1.73	726	84.80

It appears that the oldest companies, survived in average longer than more recent firms. And as well we observe the fact that they have the highest values for all the variables. The period of creation of the company is determinant in our problem. By the time, we can suppose that the market became more competitive to new entrants.

Table [7]: Properties of contractors depending on age

Contractors	Average duration (in days)	Average duration of neighbors (in days)	Average number of partners per day	Average number of contracts per day	Number of compani es	Percen tage (%)
Period 1	5393.71	2968.04	3.65	26.07	7	5.00
Period 2	2205.97	1624.58	4.05	17.45	35	25.20
Period 3	795.54	1202.63	1.92	4.16	97	69.80

We observe the same behavior for drillers; the older the companies are, the higher all the variables are. On the two previous tables, we can see that drillers and operators in proportion arrived almost together logically because a type company can only work with one of the other types. After this part and in order to find specific patterns, we create two levels for each of the three variables, number of contracts (variable A), number of partners (variable B) and duration of partners (variable C). A low level and a high level, the entities are split up into the two groups with 50% of the number each. We observe the following results.

Table [8]: Balance sheet all period operators

Configuration A B C	Number of companies	Percentage (%)	Average duration (in days)	Average age (in years)
+ + +	173	20.20	1072.00	4.00
+ + -	125	14.60	194,48	2.23
+ - +	86	10.00	731.13	3.26
+ - -	44	5.10	654.95	2.97
- - +	119	14.00	558.45	3.13
- - -	179	21.00	433.30	2.49
- + +	50	5.80	84.44	3.17
- + -	80	9.30	75.60	1.98

We are actually dealing with 856 client operators. During all the period, we can differentiate two groups whose duration is significantly lower than others (about 15.1% of the total number). These two groups have in common:

- Low number of contracts
- High number of partners

While working as an operator, it is worthless to work with many drillers when having a small number of contracts, 130 went bankrupt with these characteristics. At the opposite, operators that made it (about 30.2%) have in common:

- High number of contracts

- High duration of partners

As a consequence, the number of contracts is a really important aspect that client operators need to work on, and they should work with efficient and old drillers to increase the robustness of their network.

Table [9]: Balance sheet all period contractors

Configuration A B C	Number of companies	Percentage (%)	Average duration (in days)	Average age (in years)
+ + +	44	31.70	2254.40	6.60
+ + -	18	13.00	922.30	2.97
+ - +	8	5.70	1672.60	5.40
+ - -	0	0.00	0.00	0.00
- - +	14	10.00	578.14	2.03
- - -	47	33.80	381.91	2.49
- + +	4	2.90	1899.30	5.50
- + -	4	2.90	519.75	2.08

We do the same process for contractors, which are in a small number, solely 139. We have previously seen that they are stronger than operators. The three most successful groups of firms (about 40.3%) have in common a high duration of partners, while the two strongest ones have a high number of partners as well. The three worst (46.8%) have a low number of contracts, but it is not relevant because the second most successful group of firms has this characteristic. The two worst have a low duration of partners. To increase resilience while working as a driller, it is preferable to have a high duration of partners, working with old strong firms and avoiding recent companies. Since we have groups, we can now cross the different tables. We analyzed the values of the variables based on their duration (tables (4) and (5)), and observed the variables depending on the age of the companies (tables (6) and (7)), and we can now provide a deeper point of view. For each period, we will show the different groups for duration and even the balance sheet for the period.

Table [10-a]: Operators - period 1

Operators	Duration neighbors (in days)	Average number of partners per day	Average number of contracts per day	Number of companies	Percentage (%)
Group 1	0	0	0	0	0
Group 2	0	0	0	0	0
Group 3	0	0	0	0	0
Group 4	5288.6	1.57	10.1	5	100

Table [10-b]: Operators - period 1

Configuration A B C	Number of companies	Average duration (in days)	Percentage (%)
+ + +	5	5641.2	100
+ + -	0	0	0
+ - +	0	0	0
+ - -	0	0	0
- - +	0	0	0
- - -	0	0	0
- + +	0	0	0
- + -	0	0	0

We are starting with the oldest period and have seen that the older the company is, the longer it stands. During this time 100% of firms worked well (more than 15 years in average) with all the variables are at their high level. These operators worked with a high number of contractors that also stand for long and had a high number of contracts. We have the same profile for contractors during this period as we can see on the next figures.

Table [11-a]: Contractors - period 1

Contractors	Duration neighbors (in days)	Average number of partners per day	Average number of contracts per day	Number of companies	Percentage (%)
Group 1	0	0	0	0	0
Group 2	0	0	0	0	0
Group 3	0	0	0	0	0
Group 4	2968	3.6	26.1	7	100

Table [11-b]: Contractors - period 1

Configuration A B C	Number of companies	Average duration (in days)	Percentage (%)
+ + +	6	5617.3	86
+ + -	0	0	0
+ - +	1	4052	14
+ - -	0	0	0
- - +	0	0	0
- - -	0	0	0
- + +	0	0	0
- + -	0	0	0

We observe the same behavior for drillers, 86% have the previous profile we met, and 14% have a high number of contracts, long duration of partners but a small number of them. The average duration is lower than the previous one because of the lack of partners. This variable is meaningful here. We had only few contracts in our dataset for the first period, we may see more profiles during the next one. We study the second period, from March 2005 to October 2012, for operators first.

Table [12-a]: Operators - period 2

Operators	Duration neighbors (in days)	Average number of partners per day	Average number of contracts per day	Number of companies	Percentage (%)
Group 1	1979.8	1	1.1	16	12.8
Group 2	2236	0.8	1.8	18	14.4
Group 3	3132.3	1.3	8.5	38	30.4
Group 4	3339	1.8	10.6	53	42.4

Table [12-b]: Operators – period 2

Configuration A B C	Number of companies	Average duration (in days)	Percentage (%)
+ + +	60	1642.8	48
+ + -	2	1275	1.6
+ - +	14	1811	11.2
+ - -	8	1325.3	6.4

- - +	24	1276.8	19.2
- - -	8	958.75	6.4
- + +	9	63.3	7.2
- + -	0	0	0

Many different configurations of networks appear here, it has been a successful period for client operators. The most successful group is the one with all the variables at their highest level, companies working with numerous partners, which have working for a while and having many contracts. This group counts almost 50% of the operators that worked during this period. The group including the worst firms (7.2% of the total) worked with many partners that were resilient (high duration in average) but with only small number of contracts. This aspect is crucial for operators. In fact, this network structure has to be avoided. Now, we look at the contractors during the same period.

Table [13-a]: Contractors - period 2

Contractors	Duration neighbors (in days)	Average number of partners per day	Average number of contracts per day	Number of companies	Percentage (%)
Group 1	0	0	0	0	0
Group 2	0	0	0	0	0
Group 3	1235	2.8	7.5	6	17.1
Group 4	1710	4.3	19.5	29	82.9

Table [13-b]: Contractors - period 2

Configuration A B C	Number of companies	Average duration (in days)	Percentage (%)
+ + +	23	2260	65.6
+ + -	4	1574.5	11.4
+ - +	4	1801	11.4
+ - -	0	0	0
- - +	1	2837	2.9
- - -	1	2400	2.9
- + +	2	2794.5	5.8
- + -	0	0	0

From March 2005 to October 2012, 35 contractors appeared. About 83% worked for more than 1500 days which is a kind of resilient companies. 65.6% have all the variables at their highest level. All of them have interesting statistics. The lowest group is the one where the partners have a short duration. Another important characteristic appears, the duration of neighbors is determinant for contractors. We proceed with the most recent period.

Table [14-a]: Operators - period 3

Operators	Duration neighbors (in days)	Average number of partners per day	Average number of contracts per day	Number of companies	Percentage (%)
Group 1	1946	1	1.1	459	63.2
Group 2	2235.1	0,8	1.9	210	28.9
Group 3	2711.9	1.3	6.1	56	7.7
Group 4	5227	1.3	8.7	1	0.2

Table [14-b]: Operators - period 3

Configuration A B C	Number of companies	Average duration (in days)	Percentage (%)
+ + +	108	543.4	14.9
+ + -	123	176.9	16.9
+ - +	72	521	10
+ - -	36	506	5
- - +	95	377	13.1
- - -	171	408.7	23.6
- + +	41	89.1	5.6
- + -	80	75.6	10.9

A really impressive number of operators started working during the recent years. From November 2012 to August 2016, there are 726 of them that appeared which is more than 6 times than the previous period. Most of them went bankrupt (about 63%) and survived less than one year. The two worst group have few contracts, with a high number of partners. Only 0.2 percent of them survived more than 1500 days which is really a small part of this representative number of operators. Only one company working with strong partners with an important number of contracts.

Table [15-a]: Contractors - period 3

Contractors	Duration neighbors (in days)	Average number of partners per day	Average number of contracts per day	Number of companies	Percentage (%)
Group 1	654.1	0.9	1.1	36	37.1
Group 2	906.8	1.3	2.1	48	49.5
Group 3	1266.1	2.7	6.5	13	13.4
Group 4	0	0	0	0	0

Table [15-b]: Contractors - period 3

Configuration A B C	Number of companies	Average duration (in days)	Percentage (%)
+ + +	15	900	15.5
+ + -	14	736	14.4
+ - +	3	708.7	3.1
+ - -	0	0	0
- - +	13	404.4	13.4
- - -	46	348.3	47.3
- + +	2	1004	2.1
- + -	4	519.75	4.2

For this final period, we have 97 contractors, almost three times more than previously. Here 86.6 % survived less than 1000 days. 47.3% have worked with a small number of operators, which were not working for a long time and had few contracts. The best average duration is still in the group with the variables at their highest level. We have the following tables to summarize what we have seen previously.

Table [16]: Balance sheet of periods operators

	Number	Remarks	Structure A B C	Percentage
Period 1	5 operators	Resilient	+ + +	100%
Period 2	125 operators	Resilient	+ + +	48%
		Not resilient	- + +	7%
Period 3	726 operators	Resilient	+ + +	14.9%
		Not resilient	- + +	5.6%
		Not resilient	- + -	10.9%

Table [17]: Balance sheet of periods contractors

	Number	Remarks	Structure A B C	Percentage
Period 1	7 contractors	Resilient	+ + +	86%
		Resilient	+ - +	14%
Period 2	35 contractors	Resilient	+ + +	65%
		Not resilient	+ + -	11.4%
Period 3	97 contractors	Resilient	+ + +	15.5%
		Not resilient	- - -	47.3%
		Not resilient	- - +	13.4%

Depending on the period, several patterns appear, whereas we have a general line for both entities. We can differentiate the resilient firms from not resilient firms, we have summarized the properties on the next table.

Table [18]: Balance sheet of all the duration

	Operators	Contractors
Strong Network	High number of contracts with long duration partners	High number of partners with long duration
Not resilient network	High number of partners and small number of contracts	Short duration of partners

4. Results

As a conclusion, for our computational example, we have different properties for the two types of entities. If we consider the duration of an entity as an equivalent of the resilience of its network, we can give some properties that make a resilient network. First for contractors, the important aspect is to have an important number of partners because operators are less stable, as we have seen they are many of them and they are not active for a long time in average. Also, contractors have to be highly connected. This means that the most connections their partners have, the better it is. However, the most important part is to work with operators that have worked for a long time. Otherwise, it can lead to short duration of active contracts and possible bankruptcy. The notion of resilience for operators is quite different. The first important aspect is the number of contracts (apparently even more important than for contractors). If an operator has an important number of contracts, it is a real guarantee of success when the partners have been working for a while. However, as we have seen it is not often the case, and it can lead to bankruptcy. The worst case for an operator is to work with a high number of contractors while having a low number of contracts.

VI. Conclusions and future work

In this thesis, we provided a way to analyze corporate networks from an empirical perspective, by using a dataset with historical information regarding number and duration of contracts. The results of the proposed process provide a deep understanding of any corporate network and patterns of resilience to be applied to the structure of companies.

To improve the current methodology of this thesis, it can be of interest to include more aspects and variables into the analysis. One of the most simplifying assumptions of the research is the equivalence of the firms, which in general is not realistic. Considering firms and contracts as equal may reduce the accuracy of the research. Also, the amount of the contracts is an important variable to take into account. An important contract can in reality bring more than many small contracts, so including this variable to future research should provide accuracy.

VII. References

De Domenico, M., & Arenas, A. (2017). Modeling structure and resilience of the dark network. *Physical Review E*. https://doi.org/10.1103/PhysRevE.95.022313

Anand, K., Gai, P., Kapadia, S., Brennan, S., & Willison, M. (2013). A network model of financial system resilience. *Journal of Economic Behavior and Organization*. https://doi.org/10.1016/j.jebo.2012.04.006

Hernandez, E., & Menon, A. (2018). Acquisitions, node collapse, and network revolution. *Management Science*. https://doi.org/10.1287/mnsc.2016.2691

Colten, C. E., Hay, J., & Giancarlo, A. (2012). Community resilience and oil spills in coastal Louisiana. *Ecology and Society*. https://doi.org/10.5751/ES-05047-170305

Havlin, S. (2011). *Complex networks: Structure, robustness and function*. *Complex Networks: Structure, Robustness and Function*. https://doi.org/10.1017/CBO9780511780356

Liu, Y. Y., Slotine, J. J., & Barabási, A. L. (2011). Controllability of complex networks. *Nature*. https://doi.org/10.1038/nature10011

Squartini, T., Van Lelyveld, I., & Garlaschelli, D. (2013). Early-warning signals of topological collapse in interbank networks. *Scientific Reports*. https://doi.org/10.1038/srep03357

Schneider, C. M., Moreira, A. A., Andrade, J. S., Havlin, S., & Herrmann, H. J. (2011). Mitigation of malicious attacks on networks. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1009440108

Marnerides, A., James, C., Filho, A. S., Sait, S. Y., … Murthy, H. (2011). MULTI-LEVEL NETWORK RESILIENCE: TRAFFIC ANALYSIS, ANOMALY DETECTION AND SIMULATION. *ICTACT Journal on Communication Technology*. https://doi.org/10.21917/ijct.2011.0048

Moon, Y. H., & Jeon, Y. S. (2015). Network resilience estimation to cascading failures. In *International Conference on ICT Convergence 2015: Innovations Toward the IoT, 5G, and Smart Media Era, ICTC 2015*. https://doi.org/10.1109/ICTC.2015.7354712

Nagaishi, E., & Takemoto, K. (2018). Network resilience of mutualistic ecosystems and environmental changes: An empirical study. *Royal Society Open Science*. https://doi.org/10.1098/rsos.180706

Min, B., Yi, S. Do, Lee, K. M., & Goh, K. I. (2014). Network robustness of multiplex networks with interlayer degree correlations. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*. https://doi.org/10.1103/PhysRevE.89.042811

Kharrazi, A., Rovenskaya, E., & Fath, B. D. (2017). Network structure impacts global commodity trade growth and resilience. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0171184

Paul, G., Tanizawa, T., Havlin, S., & Stanley, H. E. (2004). Optimization of robustness of complex networks. In *European Physical Journal B*. https://doi.org/10.1140/epjb/e2004-00112-3

Sterbenz, J. P. G., Hutchison, D., Çetinkaya, E. K., Jabbar, A., Rohrer, J. P., Schöller, M., & Smith, P. (2010). Resilience and survivability in communication networks: Strategies, principles, and survey of disciplines. *Computer Networks*. https://doi.org/10.1016/j.comnet.2010.03.005

Granda, J. C., Nuño, P., Molleda, J., Usamentiaga, R., & García, D. F. (2016). Resilient overlay network for real-time interactive multimedia sessions in corporate networks. *Multimedia Systems*. https://doi.org/10.1007/s00530-015-0466-5

Quattrociocchi, W., Caldarelli, G., & Scala, A. (2014). Self-healing networks: Redundancy and structure. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0087986

Missaoui, R., Negre, E., Anggraini, D., & Vaillancourt, J. (2013). Social network restructuring after a node removal. *International Journal of Web Engineering and Technology*. https://doi.org/10.1504/IJWET.2013.052582

Gao, J., Barzel, B., & Barabási, A. L. (2016). Universal resilience patterns in complex networks. *Nature*. https://doi.org/10.1038/nature16948

Valente, A. X. C. N., Sarkar, A., & Stone, H. A. (2004). Two-Peak and Three-Peak Optimal Complex Networks. *Physical Review Letters*. https://doi.org/10.1103/PhysRevLett.92.118702

Jeude, J. A. V. L. De, Aste, T., & Caldarelli, G. (2019). The multilayer structure of corporate networks. *New Journal of Physics*. https://doi.org/10.1088/1367-2630/ab022d

INSEE, https://www.insee.fr/en/metadonnees/definition/c1927, (2016)